

Reduction principles for quantile and Bahadur-Kiefer processes of long-range dependent linear sequences

Miklós Csörgő* Rafał Kulik†

February 8, 2008

Title of the document: Reduction principles LRD
The final version (before editorial proofs) accepted to Probability Theory and
Related Fields on 15 October 2007.

Abstract

In this paper we consider quantile and Bahadur-Kiefer processes for long range dependent linear sequences. These processes, unlike in previous studies, are considered on the whole interval $(0, 1)$. As it is well-known, quantile processes can have very erratic behavior on the tails. We overcome this problem by considering these processes with appropriate weight functions. In this way we conclude strong approximations that yield some remarkable phenomena that are not shared with i.i.d. sequences, including weak convergence of the Bahadur-Kiefer processes, a different pointwise behavior of the general and uniform Bahadur-Kiefer processes, and a somewhat "strange" behavior of the general quantile process.

Keywords: long range dependence, linear processes, Bahadur-Kiefer process, quantile processes, strong approximation

Short title: Quantiles and LRD

*School of Mathematics and Statistics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6 Canada, email: mcsorgo@math.carleton.ca

†School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia, email: rkuli@maths.usyd.edu.au and Mathematical Institute, Wrocław University, Pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland.

Research supported in part by NSERC Canada Discovery Grants of Miklós Csörgő, Donald Dawson and Barbara Szyszkowicz at Carleton University

1 Introduction

Let $\{\epsilon_i, i \geq 1\}$ be a centered sequence of i.i.d. random variables. Consider the class of stationary linear processes

$$X_i = \sum_{k=0}^{\infty} c_k \epsilon_{i-k}, \quad i \geq 1. \quad (1)$$

We assume that the sequence $c_k, k \geq 0$, is regularly varying with index $-\beta$, $\beta \in (1/2, 1)$ (written as $c_k \in RV_{-\beta}$). This means that $c_k \sim k^{-\beta} L_0(k)$ as $k \rightarrow \infty$, where L_0 is slowly varying at infinity. We shall refer to all such models as long range dependent (LRD) linear processes. In particular, if the variance exists, then the covariances $\rho_k := EX_0 X_k$ decay at the hyperbolic rate, $\rho_k = L(k)k^{-(2\beta-1)} =: L(k)k^{-D}$, where $\lim_{k \rightarrow \infty} L(k)/L_0^2(k) = B(2\beta - 1, 1 - \beta)$ and $B(\cdot, \cdot)$ is the beta-function. Consequently, the covariances are not summable (cf. [14]).

Assume that X_1 has a continuous distribution function F . For $y \in (0, 1)$ define $Q(y) = \inf\{x : F(x) \geq y\} = \inf\{x : F(x) = y\}$, the corresponding (continuous) quantile function. Given the ordered sample $X_{1:n} \leq \dots \leq X_{n:n}$ of X_1, \dots, X_n , let $F_n(x) = n^{-1} \sum_{i=1}^n 1_{\{X_i \leq x\}}$ be the empirical distribution function and $Q_n(\cdot)$ be the corresponding left-continuous sample quantile function. Define $U_i = F(X_i)$ and $E_n(x) = n^{-1} \sum_{i=1}^n 1_{\{U_i \leq x\}}$, the associated uniform empirical distribution. Denote by $U_n(\cdot)$ the corresponding uniform sample quantile function.

Our purpose in this paper is to study the asymptotic behavior of sample quantiles for long range dependent sequences. This will be done in the spirit of the Bahadur-Kiefer approach (cf. [1], [16], [17]).

Assume that $E\epsilon_1^2 < \infty$. Let r be an integer and define

$$Y_{n,r} = \sum_{i=1}^n \sum_{1 \leq j_1 < \dots < j_r} \prod_{s=1}^r c_{j_s} \epsilon_{i-j_s}, \quad n \geq 1,$$

so that $Y_{n,0} = n$, and $Y_{n,1} = \sum_{i=1}^n X_i$. If $p < (2\beta - 1)^{-1}$, then

$$\sigma_{n,p}^2 := \text{Var}(Y_{n,p}) \sim n^{2-p(2\beta-1)} L_0^{2p}(n). \quad (2)$$

Define now the general empirical, the uniform empirical, the general quantile and the uniform quantile processes respectively as follows:

$$\beta_n(x) = \sigma_{n,1}^{-1} n(F_n(x) - F(x)), \quad x \in \mathbb{R}, \quad (3)$$

$$\alpha_n(y) = \sigma_{n,1}^{-1} n(E_n(y) - y), \quad y \in (0, 1), \quad (4)$$

$$q_n(y) = \sigma_{n,1}^{-1}n(Q(y) - Q_n(y)), \quad y \in (0, 1), \quad (5)$$

$$u_n(y) = \sigma_{n,1}^{-1}n(y - U_n(y)), \quad y \in (0, 1). \quad (6)$$

Assume for a while that X_i , $i \geq 1$ are i.i.d. We shall refer to this as to the *i.i.d. model*. Denote by α_n^{iid} , q_n^{iid} , u_n^{iid} the uniform empirical, general quantile, uniform quantile processes based on i.i.d. samples with the constants $\sigma_{n,1}^{-1}n$ in (4), (5), (6) replaced with \sqrt{n} . Fix $y \in (0, 1)$. Let I_y be a neighborhood of $Q(y)$ and assume that F is twice differentiable with respect to Lebesgue measure with respective first and second derivatives f and f' . Assuming that $\inf_{x \in I_y} f(x) > 0$ and $\sup_{x \in I_y} |f'(x)| < \infty$, Bahadur in [1] obtained the following *Bahadur representation* of quantiles

$$\alpha_n^{\text{iid}}(y) - f(Q(y))q_n^{\text{iid}}(y) =: R_n^{\text{iid}}(y), \quad (7)$$

with

$$R_n^{\text{iid}}(y) = O_{\text{a.s.}}(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}), \quad n \rightarrow \infty, \quad (8)$$

The process $\{R_n^{\text{iid}}(y), y \in (0, 1)\}$ is called the Bahadur-Kiefer process. Later, Kiefer proved in [16] that (8) can be strengthened to

$$R_n^{\text{iid}}(y) = O_{\text{a.s.}}(n^{-1/4}(\log \log n)^{3/4}), \quad (9)$$

which is the optimal rate. Continuing his study, in [17] Kiefer obtained the uniform version of (7), referred to later on as the *Bahadur-Kiefer representation*:

$$\sup_{y \in [0, 1]} \left| \alpha_n^{\text{iid}}(y) - f(Q(y))q_n^{\text{iid}}(y) \right| =: R_n^{\text{iid}} \quad (10)$$

where

$$R_n^{\text{iid}} = O_{\text{a.s.}}(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}), \quad n \rightarrow \infty. \quad (11)$$

Once again, the above rate is optimal. Kiefer obtained his result assuming

(K1) f has finite support and $\sup_{x \in \mathbf{R}} |f'(x)| < \infty$,

(K2) $\inf_{x \in \mathbf{R}} f(x) > 0$.

We shall refer to (K1)-(K2) as to the *Kiefer conditions*.

Further on, Csörgő and Révész [8] obtained Kiefer's result (10) under the following, weaker conditions, which shall be referred to later on as the *Csörgő-Révész conditions* (cf. also [2, Theorem 3.2.1]):

- (CsR1) f exists on (a, b) , where $a = \sup\{x : F(x) = 0\}$, $b = \inf\{x : F(x) = 1\}$,
 $-\infty \leq a < b \leq \infty$,
- (CsR2) $\inf_{x \in (a, b)} f(x) > 0$,
- (CsR3) $\sup_{x \in (a, b)} F(x)(1 - F(x)) \frac{|f'(x)|}{f^2(x)} = \sup_{y \in (0, 1)} y(1 - y) \left| \frac{f'(Q(y))}{f^2(Q(y))} \right| \leq \gamma$ with
some $\gamma > 0$,
- (CsR4) (i) $0 < A := \lim_{y \downarrow 0} f(Q(y)) < \infty$, $0 < B := \lim_{y \uparrow 1} f(Q(y)) < \infty$, or
(ii) if $A = 0$ (respectively $B = 0$) then f is nondecreasing (respectively
nonincreasing) on an interval to the right of $Q(0+)$ (respectively to the
left of $Q(1-)$).

In particular, they showed that, under (CsR1), (CsR2), (CsR3), as $n \rightarrow \infty$,

$$\sup_{n^{-1} \log \log n \leq y \leq 1 - n^{-1} \log \log n} |f(Q(y))q_n^{\text{iid}}(y) - u_n^{\text{iid}}(y)| = O_{\text{a.s.}}(n^{-1/2} \log \log n). \quad (12)$$

Additionally, if (CsR4) holds, then, as $n \rightarrow \infty$,

$$\sup_{y \in [0, 1]} |f(Q(y))q_n^{\text{iid}}(y) - u_n^{\text{iid}}(y)| = O_{\text{a.s.}}(n^{-1/2} \ell(n)). \quad (13)$$

Here, and in the sequel, $\ell(n)$ is a slowly varying function at infinity, but can be different at each place it appears (e.g. when Csörgő-Révész conditions hold, then $\ell(n) = \log \log n$). This, via the special case of (11)

$$\sup_{y \in [0, 1]} |u_n^{\text{iid}}(y) - \alpha_n^{\text{iid}}(y)| = O_{\text{a.s.}}(n^{-1/4} (\log n)^{1/2} (\log \log n)^{1/4}),$$

yields the Bahadur-Kiefer representation (11) under less restrictive conditions compared to Kiefer's assumptions. In particular, Csörgő-Révész conditions are fulfilled if F is exponential or normal. Also, if (CsR4)(i) obtains, then $\ell(n)$ in (13) is $\log \log n$. We refer to [2], [9] and [10] for more discussion of these conditions. We note in passing that taking sup over $[1/(n+1), n/(n+1)]$ instead of the whole unit interval, the statement (13) holds true assuming only the conditions (CsR1)-(CsR3) (cf. [4, Theorem 3.1], or [6, Theorem 6.3.1]).

As to LRD linear processes with partial sums $Y_{n,r}$ above, the first result on sample quantiles can be found in Ho and Hsing [15], where it is shown under Kiefer-type conditions that, as $n \rightarrow \infty$, one has for all $\beta \in (\frac{1}{2}, 1)$

$$\sup_{y \in (y_0, y_1)} |Q(y) - Q_n(y) - n^{-1} Y_{n,1}| = o_{\text{a.s.}}(n^{-(1+\lambda)} \sigma_{n,1}), \quad (14)$$

where $0 < y_0 < y_1 < 1$ are fixed and $0 < \lambda < (\beta - \frac{1}{2}) \wedge (1 - \beta)$. This means that the sample quantiles $Q_n(y)$, $y \in (y_0, y_1)$ can be approximated by the sample mean $n^{-1}Y_{n,1} = n^{-1} \sum_{i=1}^n X_i$ independently of y . This quantile process approximation is a consequence of their landmark result for empirical processes; see also [18], [22] and [23] for related studies. The best available result along these lines is due to Wu [25]. To state a particular version of his result, let F_ϵ be the distribution function of the centered i.i.d. sequence $\{\epsilon_i, i \geq 1\}$. Assume that for a given integer p , the derivatives $F_\epsilon^{(1)}, \dots, F_\epsilon^{(p+3)}$ of F_ϵ are bounded and integrable. Note that these properties are inherited by the distribution F as well (cf. [15] or [25]).

Theorem 1.1 *Let p be a positive integer. Then, as $n \rightarrow \infty$,*

$$\mathbb{E} \sup_{x \in \mathbb{R}} \left| \sum_{i=1}^n (1_{\{X_i \leq x\}} - F(x)) + \sum_{r=1}^p (-1)^{r-1} F^{(r)}(x) Y_{n,r} \right|^2 = O(\Xi_n + n(\log n)^2),$$

where

$$\Xi_n = \begin{cases} O(n), & (p+1)(2\beta-1) > 1 \\ O(n^{2-(p+1)(2\beta-1)} L_0^{2(p+1)}(n)), & (p+1)(2\beta-1) < 1 \end{cases}.$$

Using this result, under Kiefer conditions as $n \rightarrow \infty$, Wu [26] obtained

$$\sup_{y \in (y_0, y_1)} |\alpha_n(y) - f(Q(y))q_n(y) - \sigma_{n,1}^{-1} n^{-1} Y_{n,1}^2 f'(Q(y))/2| = O_{\text{a.s.}}(j_n \ell(n)), \quad (15)$$

where $j_n = n^{-(\frac{3}{4} - \frac{\beta}{2})}$ if $\beta > \frac{7}{10}$ and $j_n = n^{-(2\beta-1)}$ if $\beta \leq \frac{7}{10}$. As argued in [26, Section 7.1] this bound is sharp up to a multiplicative slowly varying function $\ell(n)$. From (15) and the central limit theorem for the partial sums $\sum_{i=1}^n X_i$ we may also deduce under Kiefer conditions and $\beta \in (\frac{1}{2}, \frac{5}{6})$, that for the Bahadur-Kiefer process

$$R_n(y) = \alpha_n(y) - f(Q(y))q_n(y) \quad (16)$$

we have weak convergence $\sigma_{n,1}^{-1} n R_n(y) \Rightarrow f'(Q(y))Z^2/2$ in $D([y_0, y_1])$, Csörgő-Révész conditions equipped with the sup-norm topology, where Z is a standard normal random variable. In particular, if $\epsilon_i, i \geq 1$ are i.i.d. standard normal random variables, then, as $n \rightarrow \infty$,

$$\sigma_{n,1}^{-1} n R_n(y) \Rightarrow \phi'(\Phi^{-1}(y))Z^2/2 \quad \text{in } D([y_0, y_1]), \quad (17)$$

where ϕ and Φ are the standard normal density and distribution functions, respectively.

This behavior is completely different compared to the i.i.d. case, for it is well known that the Bahadur-Kiefer process cannot converge weakly in the space of cadlag functions (cf., e.g., [11, Remark 2.1]).

However, this weak convergence phenomenon was first observed explicitly by Csörgő, Szyszkowicz and Wang [11] for long range dependent Gaussian sequences. For the sake of comparison with (17), assume that $\epsilon_i, i \geq 1$ are standard normal random variables and that $\sum_{k=1}^{\infty} c_k^2 = 1$. Then the X_i defined by (1) are standard normal. Define $Y_n = G(X_n)$, with some real-valued measurable function G . Let $J_l(y) = \mathbb{E} \left[\left(1_{\{F(G(X)) \leq y\}} - y \right) H_l(X) \right]$, where H_l is the l th Hermite polynomial. In particular, taking $G = F^{-1}\Phi$ we have that Y_n have the marginal distribution F . The Hermite rank is 1 and $J_1(y) = -\phi(\Phi^{-1}(y))$, and we may take $Y_n = X_n$. Note that for the Hermite rank 1, via $L(n) \sim B(2\beta - 1, 1 - \beta)L_0^2(n)$, their scaling factor $d_n^2 = n^{2-\tau D} L^\tau(n)$ (cf. (1.5) of [11]) agrees (up to a constant) with $\sigma_{n,1}^2$ of (2). Note also that $J_1(y)J_1'(y) = \phi'(\Phi^{-1}(y))$. Thus, for the uniform Bahadur-Kiefer process

$$\tilde{R}_n(y) = \alpha_n(y) - u_n(y) \quad (18)$$

we may conclude from [11, Theorem 2.3] that (see also Remark 2.22 in the present paper), as $n \rightarrow \infty$,

$$\sigma_{n,1}^{-1} n \tilde{R}_n(y) \Rightarrow \phi'(\Phi^{-1}(y)) Z^2 \quad \text{in } D([y_0, y_1]). \quad (19)$$

Comparing (17) with (19), we see that the weak limits in $D[y_0, y_1]$ of the uniform and the general Bahadur-Kiefer processes are different.

We note that Csörgő *et al.* [11] have also established the rate for the deviation of $\tilde{R}_n(y)$ from $R_n(y)$ under the Csörgő-Révész conditions. This rate, in the case of the Hermite rank 1, coincides with the scaling factor for the weak convergence of the Bahadur-Kiefer processes in (17) and (19). Since the uniform and the general Bahadur-Kiefer processes have different limits, the rate obtained for their nearness in [11] cannot be improved.

In this paper we deal with several problems. First, unlike in [15] or [26], we consider quantile and Bahadur-Kiefer processes on the whole interval $(0, 1)$ under very general conditions on the distribution function F . As it is well-known, quantile processes can have very erratic behavior on the tails. Moreover, it should be pointed out that in the LRD case, even when we deal with the associated uniform version of quantile and Bahadur-Kiefer processes, we also have to deal with the general quantile function of X_1 . We solve this problem by considering these processes with appropriate weight

functions. With this help, we can conclude various strong approximations, as well as some remarkable phenomena not shared with i.i.d. sequences, including weak convergence of the Bahadur-Kiefer processes, or different pointwise behavior of the general and uniform Bahadur-Kiefer processes. Further on, we deal with the general quantile process $q_n(y)$. Via its weak convergence, we obtain confidence intervals for the quantile function Q . Moreover, if one considers the subordinated Gaussian sequence $Y_n = G(X_n)$, then the behavior of the quantile process does not only depend on the marginals of Y_n 's and the dependence structure (i.e. the parameter β), but also on a "hidden" LRD sequence $\{X_i, i \geq 1\}$. This property cannot occur in a weakly dependent case.

Although, especially by dealing with weight functions, the paper is fairly technical, however, the choice of 'good' weight functions allow us to obtain reasonable simultaneous confidence intervals for the quantile function (see Section 2.2).

Our results are presented in Section 2. That section is concluded with a number of remarks (see Section 2.3), including a discussion of the recent paper [11]. The proofs are given in Section 3

In what follows C will denote a generic constant which may be different at each of its appearances. Also, for any sequences a_n and b_n , we write $a_n \sim b_n$ if $\lim_{n \rightarrow \infty} a_n/b_n = 1$. Further, recall that $\ell(n)$ is a slowly varying function, possibly different at each place it appears. Moreover, $f^{(k)}$ denotes the k th order derivative of f .

2 Statement of results and discussion

For discussing our results, we introduce some notation.

Let p be a positive integer and put

$$\begin{aligned} S_{n,p}(x) &= \sum_{i=1}^n (1_{\{X_i \leq x\}} - F(x)) + \sum_{r=1}^p (-1)^{r-1} F^{(r)}(x) Y_{n,r} \\ &=: \sum_{i=1}^n (1_{\{X_i \leq x\}} - F(x)) + V_{n,p}(x), \end{aligned}$$

so that $S_{n,1}(x) = nF_n(x) + f(x) \sum_{i=1}^n X_i$, and $S_{n,0}(x) = nF_n(x)$. Setting $U_i = F(X_i)$ and $x = Q(y)$ in the definition of $S_{n,p}(\cdot)$, we arrive at its

uniform version,

$$\begin{aligned}\tilde{S}_{n,p}(y) &= \sum_{i=1}^n (1_{\{U_i \leq y\}} - y) + \sum_{r=1}^p (-1)^{r-1} F^{(r)}(Q(y)) Y_{n,r} \\ &=: \sum_{i=1}^n (1_{\{U_i \leq y\}} - y) - \tilde{V}_{n,p}(y).\end{aligned}\tag{20}$$

Recall that

$$R_n(y) = \alpha_n(y) - f(Q(y))q_n(y), \quad y \in (0, 1),$$

is the Bahadur-Kiefer process and

$$\tilde{R}_n(y) = \alpha_n(y) - u_n(y), \quad y \in (0, 1),$$

is the uniform Bahadur-Kiefer process.

We shall consider the following assumptions on the distribution function F .

(A(p)) The functions $(f^{(r-1)} \circ Q)^{(1)}(y)$, $r = 1, \dots, p$, are uniformly bounded. The integer p will be chosen appropriately in the sequel.

(B) The function $(f \circ Q)^{(2)}(y)$ is uniformly bounded.

(C(p)) For $r = 0, \dots, p-1$,

$$\sup_{y \in (0,1)} \frac{f^{(r+1)}(Q(y))}{f(Q(y))} (y(1-y))^{1/2} = O(1).$$

2.1 Strong approximations

Let

$$\begin{aligned}a_n &= \sigma_{n,1} n^{-1} \log \log n = n^{-(\beta-\frac{1}{2})} L_0(n) \log \log n, \\ b_n &= \sigma_{n,1}^2 n^{-1} a_n (\log \log n)^{1/2} = n^{-(3\beta-\frac{5}{2})} L_0^3(n) (\log \log n)^{3/2}, \\ c_n &= \sigma_{n,1}^{-1} b_n (\log n)^{1/2} = n^{-(2\beta-1)} L_0^2(n) (\log \log n)^{3/2} (\log n)^{1/2}, \\ d_{n,p} &= \begin{cases} n^{-(1-\beta)} L_0^{-1}(n) (\log n)^{5/2} (\log \log n)^{3/4}, & (p+1)(2\beta-1) > 1 \\ n^{-p(\beta-\frac{1}{2})} L_0^p(n) (\log n)^{1/2} (\log \log n)^{3/4}, & (p+1)(2\beta-1) < 1 \end{cases}, \\ b_{n,p} &= \sigma_{n,1}^2 n^{-1} d_{n,p} (\log \log n)^{1/2},\end{aligned}$$

and

$$\delta_n = n^{-(2\beta-1)} L_0^2(n) (\log \log n).$$

2.1.1 Reduction principles for the uniform quantile process

First, we deal with reduction principles for quantiles. Ho and Hsing, [15, p. 1003] asked, whether there was an expansion for the quantile process which mirrors that in their Theorem 2.1 for the empirical process. We have the following result.

Theorem 2.1 *Assume (B), and either (A(1)) or (A(2)) according to $\beta \geq 3/4$ or $\beta < 3/4$. Then, under the conditions of Theorem 1.1, as $n \rightarrow \infty$, we have,*

$$\sup_{y \in (0,1)} \left| u_n(y) + \sigma_{n,1}^{-1} f(Q(y)) \sum_{i=1}^n X_i \right| = \begin{cases} O_{a.s.}(d_{n,1}), & \text{if } \beta \geq 3/4, \\ O_{a.s.}(a_n), & \text{if } \beta < 3/4. \end{cases} \quad (21)$$

If $\beta < \frac{3}{4}$, the bound is optimal.

To remove assumptions (A) and (B) we shall consider a (possibly) *weighted* approximation of the uniform quantiles. Define $\psi_1(y)$ in the following way. If $\beta < \frac{3}{4}$, then $\psi_1(y) = 1$ if (C(2)) holds, and $\psi_1(y) = (y(1-y))^{\gamma-\frac{1}{2}+\mu}$, $\mu > 0$ otherwise. If $\beta \geq \frac{3}{4}$, $\psi_1(y) = (y(1-y))^{\gamma+\mu}$.

Theorem 2.2 *Let $p = 2$. Then, under the conditions of Theorem 1.1, as $n \rightarrow \infty$, we have,*

$$\sup_{y \in (0,1)} \psi_1(y) \left| u_n(y) + \sigma_{n,1}^{-1} f(Q(y)) \sum_{i=1}^n X_i \right| = \begin{cases} O_{a.s.}(d_{n,1}), & \text{if } \beta \geq 3/4, \\ O_{a.s.}(a_n), & \text{if } \beta < 3/4. \end{cases} \quad (22)$$

If $\beta < \frac{3}{4}$, the bound is optimal.

From Theorems 2.1 or 2.2 and Lemma 3.6 below we have the following reduction principle for quantiles, which mirrors that for the empirical process. In order to state the result, redefine ψ_1 to be 1 if, for a given p , (A(p)) holds, and be it as before otherwise. The result is stated for $\beta < \frac{3}{4}$, only in order to avoid the additional term coming from $d_{n,1}$.

Corollary 2.3 $\beta < \frac{3}{4}$. *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. Assume that either (A(p)) and (B), or (C(p)) hold. Under the conditions of Theorem 1.1, as $n \rightarrow \infty$,*

$$\begin{aligned} & \sup_{y \in (0,1)} \sigma_{n,p}^{-1} \psi_1(y) |y - U_n(y) + n^{-1} \tilde{V}_{n,p}(y)| \\ &= O_{a.s.}(n^{-(2\beta-p\frac{2\beta-1}{2})} L_0^{2-p}(n) \log \log n (\log n)^{1/2}). \end{aligned}$$

2.1.2 Approximations of the uniform Bahadur-Kiefer process

Similarly to the uniform quantile process, in Theorem 2.4 we obtain strong approximation of the uniform Bahadur-Kiefer process on the whole interval $(0, 1)$ on assuming (A) and (B).

Theorem 2.4 *Assume (B), and either (A(2)) or (A(3)) according to $\beta \geq 2/3$ or $\beta < 2/3$. Under the conditions of Theorem 1.1, as $n \rightarrow \infty$,*

$$\sup_{y \in (0,1)} \left| \tilde{R}_n(y) - n^{-1} \sigma_{n,1}^{-1} f^{(1)}(Q(y)) \left(\sum_{i=1}^n X_i \right)^2 \right| = \begin{cases} O_{a.s.}(d_{n,2}), & \text{if } \beta \geq 2/3, \\ O_{a.s.}(c_n), & \text{if } \beta < 2/3. \end{cases} \quad (23)$$

To remove assumptions (A) and (B), we shall consider a *weighted* approximation of the uniform quantile and Bahadur-Kiefer processes. Define for arbitrary $\mu > 0$,

$$\psi_2(y) = \begin{cases} (y(1-y))^{1+\mu}, & \text{if } \beta < \frac{3}{4} \text{ and } (C(3)); \\ (y(1-y))^{1+\mu}, & \text{if } \beta < \frac{3}{4}, \gamma < \frac{3}{2} \text{ and not } (C(3)); \\ (y(1-y))^{\gamma-\frac{1}{2}+\mu}, & \text{if } \beta < \frac{3}{4} \text{ and } \gamma \geq \frac{3}{2}; \\ (y(1-y))^{\gamma+\mu}, & \text{if } \beta \geq \frac{3}{4}. \end{cases}$$

Theorem 2.5 *Under the conditions of Theorem 1.1, as $n \rightarrow \infty$,*

$$\sup_{y \in (0,1)} \psi_2(y) \left| \tilde{R}_n(y) - n^{-1} \sigma_{n,1}^{-1} f^{(1)}(Q(y)) \left(\sum_{i=1}^n X_i \right)^2 \right| = \begin{cases} O_{a.s.}(d_{n,2}), & \text{if } \beta \geq 2/3, \\ O_{a.s.}(c_n), & \text{if } \beta < 2/3. \end{cases} .$$

From Theorem 2.4 or 2.5 and Lemma 3.6 below, we obtain the reduction principle for the distance between the uniform empirical and the uniform quantile processes, similar to that of Corollary 2.3. Further, an immediate corollary to Theorem 2.4, via the LIL for partial sums $\sum_{i=1}^n X_i$ (see (32) below), is the following result.

Corollary 2.6 *Under the conditions of Theorem 2.4, if $\beta < \frac{3}{4}$,*

$$\limsup_{n \rightarrow \infty} \sigma_{n,1}^{-1} n (\log \log n)^{-1} \sup_{y \in (0,1)} |\tilde{R}_n(y)| \stackrel{a.s.}{=} c(\beta, 1) \sup_{y \in (0,1)} |f^{(1)}(Q(y))|, \quad (24)$$

where $c^2(\beta, p) = \left(\int_0^\infty x^{-\beta} (1+x)^{-\beta} dx \right) (1-\beta)^{-1} (3-2\beta)^{-1}$.

Corollary 2.7 *Under the conditions of Theorem 2.4, if $\beta < \frac{3}{4}$,*

$$\sigma_{n,1}^{-1} n \tilde{R}_n(y) \Rightarrow f^{(1)}(Q(y)) Z^2.$$

The corresponding results can also be stated in the setting of Theorem 2.5.

2.1.3 Approximation of the general Bahadur-Kiefer process

As for the general Bahadur-Kiefer process, a typical approach in the i.i.d. case is to approximate the normalized quantiles $f(Q(y))q_n(y)$ via the uniform quantiles and then use this to generalize all results valid in the uniform case to the general one, as described in the Introduction (cf. (12), (13)). This approach was also followed in [11, Section 4] as well. However, this cannot work in the LRD case, for then the uniform and general Bahadur-Kiefer processes have different limits (cf. (17), (19)). Moreover, assumptions (A) and (B) do not help in this case.

With arbitrary $\mu > 0$, define

$$\psi_3(y) = \begin{cases} (y(1-y))^{1+\mu}, & \text{if } \beta < \frac{3}{4} \text{ and } (C(3)); \\ (y(1-y))^{2\gamma-1+\mu}, & \text{if } \beta < \frac{3}{4}, \text{ and not } (C(3)); \\ (y(1-y))^{2+2\gamma+\mu}, & \text{if } \beta \geq \frac{3}{4}. \end{cases}$$

We have the following result.

Theorem 2.8 *Under the conditions of Theorem 1.1 we have with some $C_0 > 0$, as $n \rightarrow \infty$,*

$$\begin{aligned} & \sup_{y \in (C_0\delta_n, 1-C_0\delta_n)} \psi_3(y) \left| R_n(y) - n^{-1}\sigma_{n,1}^{-1} \frac{f^{(1)}(Q(y))}{2} \left(\sum_{i=1}^n X_i \right)^2 \right| \\ &= \begin{cases} O_{a.s.}(d_{n,2}), & \text{if } \beta \geq 2/3, \\ O_{a.s.}(c_n), & \text{if } \beta < 2/3. \end{cases} \end{aligned} \quad (25)$$

If $\gamma = 1$ then the above estimate is valid on $(0, 1)$.

The (weighted) almost sure behavior of $R_n(\cdot)$ and (weighted) convergence can be obtained in the same way as that of $\tilde{R}_n(\cdot)$ in Corollaries 2.6 and 2.7.

2.2 Weak behavior of the general quantile process and its consequences

Ho and Hsing's result (14) would suggest that it should be possible to approximate $q_n(y)$ at least on the expanding intervals, $(n^{-1}, 1-n^{-1})$. However, as we will explain below, this is not the case.

Let $\psi_4(y) = 1$ or $y(1-y)$ according to $\beta < \frac{3}{4}$ or $\beta \geq \frac{3}{4}$, respectively.

Proposition 2.9 *Assume (CsR1)-(CsR4). Then*

$$\sup_{y \in (0,1)} \psi_4(y) |f(Q(y))q_n(y) - u_n(y)| = O_a(\sigma_{n,1} n^{-1} \ell(n)), \quad (26)$$

where $O_a = O_{a.s.}$ if $\gamma = 1$, and $O_a = O_P$ if $\gamma > 1$.

Corollary 2.10 Assume (CsR1)-(CsR4). Then, under the conditions of either Theorem 2.1 or 2.2, as $n \rightarrow \infty$,

$$\sup_{y \in (0,1)} \psi_1(y) f(Q(y)) \left| q_n(y) + \sigma_{n,1}^{-1} \sum_{i=1}^n X_i \right| = O_P(n^{-(\beta-\frac{1}{2})} \ell(n)).$$

Corollary 2.11 Assume (CsR1)-(CsR4). Then, under the conditions of either Theorem 2.1 or 2.2, as $n \rightarrow \infty$,

$$\sup_{y \in (n^{-1}, 1-n^{-1})} (y(1-y))^\nu \left| q_n(y) + \sigma_{n,1}^{-1} \sum_{i=1}^n X_i \right| = o_P(1),$$

where

$$\begin{aligned} \nu &> \gamma - (\beta - \tfrac{1}{2}), & \text{if } \beta < \tfrac{3}{4} \text{ and either (A(2)) or (C(2));} \\ \nu &> 2\gamma - \beta, & \text{if } \beta < \tfrac{3}{4} \text{ and neither (A(2)) nor (C(2));} \\ \nu &> 2\gamma - (\beta - \tfrac{1}{2}), & \text{if } \beta \geq \tfrac{3}{4}. \end{aligned}$$

From this result one obtains the following simultaneous confidence bounds, which cover all the data available for $y \in (n^{-1}, 1-n^{-1})$,

$$Q_n(y) - \sigma_{n,1} n^{-1} c_\nu z_\alpha (y(1-y))^{-\nu} \leq Q(y) \leq Q_n(y) + \sigma_{n,1} n^{-1} c_\nu z_\alpha (y(1-y))^{-\nu},$$

where z_α is the $(1-\alpha/2)$ -quantile of the standard normal law, and

$$c_\nu = \sup_{y \in (0,1)} (y(1-y))^\nu.$$

Another consequence of Corollary 2.10 is that for some $k_n = k_n(\gamma, \beta) \rightarrow 0$, as $n \rightarrow \infty$,

$$\sup_{y \in (k_n, 1-k_n)} \left| q_n(y) + \sigma_{n,1}^{-1} \sum_{i=1}^n X_i \right| = o_P(1),$$

and thus

$$q_n(y) 1_{\{y \in (k_n, 1-k_n)\}} \Rightarrow Z. \quad (27)$$

Optimally, one would hope to obtain weak convergence on $(n^{-1}, 1-n^{-1})$, but this is not a good way to treat quantiles in the LRD case at all. To see this, recall the subordinated Gaussian model $Y_n = G(X_n)$. Take $G = F^{-1}\Phi$. For the uniform sample quantile process $u_n(y)$ associated with the sequence $\{Y_n, n \geq 1\}$ one obtains in the spirit of [11, Proposition 2.2] (see [7] for a correct proof)

$$\sup_{y \in (0,1)} \left| u_n(y) + \sigma_{n,1}^{-1} \phi(\Phi^{-1}(y)) \sum_{i=1}^n X_i \right| = O_P(n^{-(\beta-\frac{1}{2})} \ell(n)). \quad (28)$$

Moreover, from [11, Proposition 4.2], if the distribution F of $Y = G(X)$ fulfills (CsR1)-(CsR3), then for some $k'_n \rightarrow 0$,

$$\sup_{y \in (k'_n, 1-k'_n)} \left| f(Q(y))q_n(y) + \sigma_{n,1}^{-1} \phi(\Phi^{-1}(y)) \sum_{i=1}^n X_i \right| = O_P(n^{-(\beta-\frac{1}{2})} \ell(n)), \quad (29)$$

where $q_n(y)$ is the general quantile process associated with Y_n . Thus,

$$q_n(y)1_{\{y \in (k'_n, 1-k'_n)\}} \Rightarrow \frac{\phi(\Phi^{-1}(y))}{f(Q(y))} Z, \quad (30)$$

provided $\frac{\phi(\Phi^{-1}(y))}{f(Q(y))}$ is uniformly bounded. In particular, if f is exponential, then this is not the case. Consequently, we may have two LRD models, both with the same covariance structure, both with the same exponential marginals, say, so that in case of (1) the general quantile process converges, while in the subordinated Gaussian case it does not converge (cf. (27) and (30), respectively). On the other hand, in both cases, the empirical processes have normal limits scaled by a deterministic function. In other words, subordination can completely change convergence properties of quantile processes, even if the empirical processes behave in the same way in the subordinated and non-subordinated cases. The weight function $(y(1-y))^\nu$ solves this problem somehow.

2.2.1 Trimmed means

In the model (1), assume that X_i are symmetric. From (27) one easily obtains

$$\sigma_{n,1}^{-1} \left| \sum_{i=[nk_n]}^{[n(1-k_n)]} X_i \right| = \left| \int_{k_n}^{1-k_n} q_n(y) dy \right| \xrightarrow{d} |Z|.$$

On the other hand, since $EX_1 = 0$, $\left| \int_0^1 q_n(y) dy \right| = \sigma_{n,1}^{-1} |\sum_{i=1}^n X_i| \xrightarrow{d} |Z|$. If $k_n < l_n \rightarrow 0$ then the result remains true by considering weak convergence in (27) on $(l_n, 1-l_n)$ and then arguing as in the case of k_n . Summarizing,

Corollary 2.12 *Assume (CsR1)-(CsR4) and that X_i are symmetric. Let $k_n \leq l_n \rightarrow 0$. Then, under the conditions of either Theorem 2.1 or 2.2,*

$$\sigma_{n,1}^{-1} \sum_{i=[nl_n]}^{[n(1-l_n)]} X_i \xrightarrow{d} Z. \quad (31)$$

The result (31) states essentially that, whatever trimming we consider, the deleted part is negligible.

However, it should be mentioned that this approach to the trimmed sums is not the optimal one. The problem is considered in more details in [19] and [20] via studying integral functionals of the empirical process (see e.g. [5] for the description of the method in the i.i.d case).

2.3 Remarks

We start with pointing out some phenomena which are exclusive for LRD sequences.

Remark 2.13 As mentioned in the Introduction, it was observed explicitly in [11] and can be concluded from [26] that the uniform Bahadur-Kiefer process (in case of [11]) and, under appropriate conditions, the general Bahadur-Kiefer process ([26]) converge in $D([y_0, y_1])$ for a particular choice of the parameter β . From our results we conclude that both processes converge weakly in $D([0, 1])$ if $\beta < \frac{3}{4}$. This is striking difference compared to the i.i.d. case, for in the latter case these processes cannot converge weakly (cf. [16], [17]). Considering pointwise convergence, in the i.i.d. case the uniform and the general Bahadur-Kiefer processes converge to the same limit (cf. [10] for a review). Here, the pointwise limits are different, on account of different weak limits.

Remark 2.14 Unlike in the i.i.d case, to study the distance between the uniform empirical and the uniform quantile processes, we need to control the general quantile process, which can be done via controlling the quantile and density quantile functions associated with X_i . The reason for this is that the uniform quantile process contains information regarding the marginal behavior of random variables X_i . This is visible from Theorems 2.1 and 2.2 - the uniform quantile process depends on the density-quantile function $f(Q(y))$ associated with X_1 . As can be seen in (28), this remains true in the subordinated case $Y_i = G(X_i)$ as well, namely the uniform quantile process contains information about the marginals of X_i , not of Y_i . This has an impact on the behavior of general quantiles, as described in Section 2.2.

We continue with some technical remarks concerning assumptions and results above.

Remark 2.15 We comment on the different rates in our theorems, according to different choices of β .

If $p = 1$ then $a_n = o(d_{n,1})$, if $p = 2$ (so that $\beta < 3/4$), then $d_{n,2} = o(a_n)$, and then optimal rates are attained in Theorems 2.1 and 2.2. Taking higher order expansions ($p \geq 3$) does not improve rates and requires additional restrictions on β and conditions on F , either (A(p)) or (C(p)).

Likewise, if $p = 1, 2$, then $c_n = o(d_{n,p})$. If $p = 3$ ($\beta < \frac{2}{3}$), then $d_{n,3} = o(c_n)$. Then we can identify (but not prove !!) optimal rates in Theorems 2.4, 2.5. We conjecture, that the bound in Theorem 2.4 (at least for $\beta < \frac{2}{3}$) is valid without the $(\log n)^{1/2}$ term due to the following conjecture.

Conjecture 1 *For any $p \geq 1$,*

$$\limsup_{n \rightarrow \infty} \sigma_{n,p}^{-1} (\log \log n)^{-p/2} Y_{n,p} \stackrel{\text{a.s.}}{=} c(\beta, p),$$

where $c(\beta, p)$ is as in Corollary 2.6.

Further, on comparing Theorem 2.8 with (15) we can see that the method in [26] leads to better rates for β close to 1. We lose some rates for β close to 1, since then the error in the reduction principle dominates. On the other hand, Wu's method is unlikely to work when one wants to deal with approximations on the whole interval $(0, 1)$, which was our main goal. In fact, in view of a weighted law of the iterated logarithm (see Lemma 3.10), it is not likely that in the case $\beta \geq \frac{3}{4}$ the estimates on $(0, 1)$ can be obtained with optimal rates, unless the rate $d_{n,p}$ is improved.

Remark 2.16 Wu in his paper [25] has in fact some weaker conditions on F_ϵ , than those stated in Theorem 1.1. Also, here, we avoid the boundary case $(p+1)(2\beta-1) = 1$. Furthermore, under stronger regularity conditions on the distribution of ϵ_1 , the reduction principle (with worse rates) for the empirical process remains true provided $E|\epsilon|^{2+\delta} < \infty$, $\delta > 0$ (see [13]). Thus, some of the results here remain valid under the Giraitis and Surgailis conditions in [13]. However, to prove Theorems 2.2 and 2.5 we require Lemma 3.9 below, where the rates in the reduction principle for Theorem 1.1 are crucial.

Remark 2.17 We comment on assumptions (A(p)), (B) and (C(p)) on the distribution function F . Note that $-(f \circ Q)^{(1)}(y) = J(y)$ is the so-called score function (cf. e.g. [2, p. 7]), thus (A(1)) requires uniform boundness of the latter. This is not valid if one takes the standard normal distribution for example. The assumptions (A(p)), $p \geq 1$ are fulfilled if one takes the exponential, logistic, or Pareto distribution $f(x) = \alpha(x^{1+\alpha})^{-1}$, $x > 1$, $\alpha > 0$. Assumption (B) is fulfilled if one takes exponential, logistic, or Pareto with $\alpha > 1$. The latter constrain $\alpha > 1$ is relevant, since in view of Theorem 1.1 we work under the condition $E\epsilon^4 < \infty$ and, consequently, $EX^4 < \infty$.

Further, $(C(p))$, $p \geq 1$, is fulfilled in the Pareto case and for the standard normal case. Thus, essentially, most of the "practical" parametric families fulfill either $(A(p))$ or $(C(p))$.

Further, in the LRD case (1) it is very unlikely that f has bounded support (from either side). Moreover, to use of Theorem 1.1, we need $E\epsilon = 0$ and $f_\epsilon = F'_\epsilon$ to be smooth. Consequently, the same properties are transferred to X and its density f . Therefore, to make use Theorem 1.1 and assumptions $(A(p))$ and (B) simultaneously, we should consider the above comments for double exponential or symmetric Pareto, appropriately smoothed around the origin. Nevertheless, the main issue of assumptions $(A(p))$, (B) and $(C(p))$ is the tail behavior.

Remark 2.18 As for the general quantile and the general Bahadur-Kiefer processes, in order to obtain their approximations on the whole interval, we assumed the monotonicity property (CsR4). In principle, as in the i.i.d. case, (cf. [4]), it should be possible to obtain their approximations on the "practical" interval $(n^{-1}, 1 - n^{-1})$ without (CsR4).

Remark 2.19 We now discuss the weights which appear in our theorems. As mentioned in Remark 2.14, the LRD sequences based uniform quantile process "feels" the general quantile function. In the i.i.d. case one knows that for $\mu > 0$

$$\limsup_{n \rightarrow \infty} \sup_{y \in (0,1)} (y(1-y))^\mu |Q(y) - Q_n^{\text{iid}}(y)| < \infty$$

almost surely if and only if $\int_{-\infty}^{\infty} |u|^{1/\mu} dF(u) < \infty$ (see [2, p. 98] for a tribute to David Mason in this regard). Therefore, our weight functions $(y(1-y))^\kappa$, with some $\kappa > 0$, appear to be natural to use.

We also note that instead of the weight $(y(1-y))^{1+\kappa}$, $\kappa > 0$, we may consider $f^{\kappa'}(Q(y))$ as a weight function, where κ' depends on both κ and γ .

Remark 2.20 In Theorem 2.8, in case $\gamma > 1$, the approximation *in probability* remains valid on $(0, 1)$ (see also Proposition 2.9). We are not able to do this almost surely, since we do not have a precise knowledge about the LRD behavior of order statistics (see the proof of Proposition 2.9).

Remark 2.21 The bound in Theorem 2.1 is determined by the behavior of the Bahadur-Kiefer process $\tilde{R}_n(y)$ (compare Theorem 2.1 with (24)). This is somehow similar to the i.i.d. case. One knows that on an appropriate probability space, $\sup_{y \in (0,1)} |\alpha_n^{\text{iid}}(y) - B_n(y)| = O_{\text{a.s.}}(n^{-1/2} \log n)$,

where $B_n(\cdot)$ are appropriate Brownian bridges. Further, via (10) we can see that with the same Brownian bridges we have $\sup_{y \in (0,1)} |u_n^{\text{iid}}(y) - B_n(y)| = O_{\text{a.s.}}(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4})$. We may for example refer to [9] and [10] for more details.

Remark 2.22 Recall, from Section 2.2, our lines the subordinated Gaussian case $Y = G(X)$. We have $J_1(y) = -\phi(\Phi^{-1}(y))$, where ϕ, Φ are the standard normal density and distribution function. Csörgő, Szyszkowicz and Wang in [11] proved their Proposition 2.2 assuming (cf. also their Remark 2.1) their Assumption A. However, what is really used in their proof is that J_1 has, in particular, uniformly bounded first order derivative, which is not true, since $J_1'(y) = -\Phi^{-1}(y)$. Consequently, their Proposition 2.2 and all its consequences in their Sections 2.1 and 2.2 are valid only if one restricts them to intervals $[y_0, y_1]$, or assumes that $Y = G(X)$ has finite support. This actually is the reason that we considered assumptions (A(p)), (B) and/or weighted approximations. Clearly, the non-subordinated Gaussian case can be treated as in the setting of Theorems 2.2, 2.5 and 2.8 with $\gamma = 1$ (recall that (C(p)) holds in the Gaussian case). For the general treatment we refer to [7].

Also, as noted already in our Section 2.1.3, results for the general Bahadur-Kiefer process cannot be concluded from an approximation of the latter by the uniform one. Hence, the proposed proofs for Theorems 4.1, 4.2 of [11] via the invariance principle of Proposition 4.2 cannot work and, in view of [26], the claimed limiting processes can at best be correct if multiplied by $1/2$.

In Section 3 of [11] the authors consider $V_n(t) = 2\sigma_{n,1}^{-1}n \int_0^t \tilde{R}_n(y)dy$ and $Q_n(t) = V_n(t) - \alpha_n^2(t)$, the so-called uniform Vervaat and Vervaat Error processes. As a consequence of our comments so far on paper [11], we note that the results in this section are valid only if $G(X)$ has finite support. An extension is possible if one has assumptions like (A(p)) and (B). This, however, is out of the scope of this paper.

3 Proofs

3.1 Preliminary results

We recall the following law of the iterated logarithm for partial sums $\sum_{i=1}^n X_i$ (see, e.g., [24]):

$$\limsup_{n \rightarrow \infty} \sigma_{n,1}^{-1} (\log \log n)^{-1/2} \left| \sum_{i=1}^n X_i \right| \stackrel{\text{a.s.}}{=} c(\beta, 1), \quad (32)$$

where $c(\beta, 1)$ is defined in Corollary 2.6.

Lemma 3.1 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. Then, as $n \rightarrow \infty$,*

$$Y_{n,p} = O_{\text{a.s.}}(\sigma_{n,p}(\log n)^{1/2} \log \log n). \quad (33)$$

Proof. Let $B_n^2 = \sigma_{n,p}^2 \log n (\log \log n)^2$. By (2), [26, Lemma 4] and Karata's Theorem we have

$$\begin{aligned} \left\| \frac{Y_{n,p}}{B_{2^d}} \right\|_2^2 &\leq \frac{1}{B_{2^d}} \left(\sum_{j=0}^d 2^{(d-j)/2} \sigma_{2^j,p} \right)^2 \leq \frac{2^d}{B_{2^d}} \left(\sum_{j=0}^d 2^{j(1-p(2\beta-1))/2} L_0^p(2^j) \right)^2 \\ &\sim \frac{2^d}{B_{2^d}} 2^{2d-dp(2\beta-1)} L_0^{2p}(2^d) \sim d^{-1} (\log d)^{-2}. \end{aligned}$$

Therefore, the result follows by the Borel-Cantelli lemma. ◉

As an easy consequence of (32) and (33) we obtain the next result.

Lemma 3.2 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. We have*

$$\limsup_{n \rightarrow \infty} \sigma_{n,1}^{-1} (\log \log n)^{-1/2} \sup_{y \in (0,1)} |\tilde{V}_{n,p}(y)| \stackrel{\text{a.s.}}{=} c(\beta, 1). \quad (34)$$

Using Theorem 1.1 and the same argument as in the proof of Lemma 3.1, we obtain

$$\begin{aligned} &\sigma_{n,p}^{-1} \sup_{x \in \mathbf{R}} |S_{n,p}(x)| \\ &= \begin{cases} O_{\text{a.s.}}(n^{-(\frac{1}{2}-p(\beta-\frac{1}{2}))} L_0^{-p}(n) (\log n)^{5/2} (\log \log n)^{3/4}), & (p+1)(2\beta-1) > 1 \\ O_{\text{a.s.}}(n^{-(\beta-\frac{1}{2})} L_0(n) (\log n)^{1/2} (\log \log n)^{3/4}), & (p+1)(2\beta-1) < 1 \end{cases}. \end{aligned}$$

Since (see (2))

$$\frac{\sigma_{n,p}}{\sigma_{n,1}} \sim n^{-(\beta-\frac{1}{2})(p-1)} L_0^{p-1}(n), \quad (35)$$

we obtain

$$\begin{aligned} & \sup_{x \in \mathbf{R}} |\beta_n(x) + \sigma_{n,1}^{-1} V_{n,p}(x)| = \\ &= \frac{\sigma_{n,p}}{\sigma_{n,1}} \sup_{x \in \mathbf{R}} \left| \sigma_{n,p}^{-1} \sum_{i=1}^n (1_{\{X_i \leq x\}} - F(x)) + \sigma_{n,p}^{-1} V_{n,p}(x) \right| = o_{\text{a.s.}}(d_{n,p}). \end{aligned}$$

Consequently, via $\{\alpha_n(y), y \in (0, 1)\} = \{\beta_n(Q(y)), y \in (0, 1)\}$,

$$\sup_{y \in (0,1)} |\alpha_n(y) + \sigma_{n,1}^{-1} \tilde{V}_{n,p}(y)| = O_{\text{a.s.}}(d_{n,p}). \quad (36)$$

Remark 3.3 For convenient reference, we collect here various relations between constants. Recall that $d_{n,2} = o(a_n)$ provided $\beta < \frac{3}{4}$, and $d_{n,3} = o(c_n)$, provided $\beta < \frac{2}{3}$. Further, $\sigma_{n,1}^{-1} b_{n,p} = o(d_{n,p})$. It is not necessarily true that $\sigma_{n,1}^{-1} = o(d_{n,p})$, but it is always true that $\sigma_{n,1}^{-1} = o(a_n)$.

3.2 Proof of Theorems 2.1 and 2.4

First, we bound the distance between the uniform empirical and uniform quantile processes.

Lemma 3.4 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. Assume (A(p)). Under the conditions of Theorem 1.1 we have, as $n \rightarrow \infty$,*

$$\sup_{y \in (0,1)} |u_n(y) - \alpha_n(y)| = O_{\text{a.s.}}(a_n) + O_{\text{a.s.}}(d_{n,p}).$$

Proof. Note that

$$\begin{aligned} u_n(y) &= \sigma_{n,1}^{-1} n(E_n(U_n(y)) - U_n(y)) - \sigma_{n,1}^{-1} n(E_n(U_n(y)) - y) \\ &= \sigma_{n,1}^{-1} n(E_n(U_n(y)) - U_n(y)) + O_{\text{a.s.}}(\sigma_{n,1}^{-1}) = \alpha_n(U_n(y)) + O(\sigma_{n,1}^{-1}). \end{aligned} \quad (37)$$

Thus, by (36),

$$\begin{aligned} & \sup_{y \in (0,1)} |u_n(y) - \alpha_n(y)| \\ &= \sup_{y \in (0,1)} |\alpha_n(U_n(y)) - \alpha_n(y)| + O_{\text{a.s.}}(\sigma_{n,1}^{-1}) \\ &\leq \sigma_{n,1}^{-1} \sup_{y \in (0,1)} |\tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y))| + O_{\text{a.s.}}(\sigma_{n,1}^{-1}) + O_{\text{a.s.}}(d_{n,p}). \end{aligned} \quad (38)$$

Accordingly, in view of Assumptions (A(p)), (B), we have to control

$$\sup_{y \in (0,1)} |f(Q(y)) - f(Q(U_n(y)))| \left| \sum_{i=1}^n X_i \right| \leq C \sup_{y \in (0,1)} |y - U_n(y)| \left| \sum_{i=1}^n X_i \right| \quad (39)$$

and

$$\begin{aligned} & \sup_{y \in (0,1)} \sum_{r=2}^p \left| f^{(r-1)}(Q(y)) - f^{(r-1)}(Q(U_n(y))) \right| |Y_{n,r}| \\ & \leq C \sup_{y \in (0,1)} |y - U_n(y)| \left| \sum_{r=2}^p Y_{n,r} \right|. \end{aligned} \quad (40)$$

From (34) and (36) one obtains

$$\limsup_{n \rightarrow \infty} (\log \log n)^{1/2} \sup_{y \in (0,1)} |\alpha_n(y)| \stackrel{\text{a.s.}}{=} c(\beta, 1).$$

Consequently, as $n \rightarrow \infty$,

$$\begin{aligned} \sup_{y \in (0,1)} |y - U_n(y)| &= \sup_{y \in (0,1)} \sigma_{n,1} n^{-1} |u_n(y)| = \sup_{y \in (0,1)} \sigma_{n,1} n^{-1} |\alpha_n(y)| \\ &= O_{\text{a.s.}}(\sigma_{n,1} n^{-1} (\log \log n)^{1/2}) = O_{\text{a.s.}}(a_n). \end{aligned} \quad (41)$$

Therefore, on combining (32), (39), (41), as $n \rightarrow \infty$, one obtains

$$\sup_{y \in (0,1)} \sigma_{n,1}^{-1} |f(Q(y)) - f(Q(U_n(y)))| \left| \sum_{i=1}^n X_i \right| = O_{\text{a.s.}}(a_n). \quad (42)$$

Having (33), (40) and (42), as $n \rightarrow \infty$, we conclude

$$\sup_{y \in (0,1)} \sigma_{n,1}^{-1} |\tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y))| = O_{\text{a.s.}}(a_n). \quad (43)$$

Thus, by (38) and (43), as $n \rightarrow \infty$,

$$\sup_{y \in (0,1)} |u_n(y) - \alpha_n(y)| = O_{\text{a.s.}}(a_n) + O(\sigma_{n,1}^{-1}) + O_{\text{a.s.}}(d_{n,p}),$$

and hence the result follows. ◉

If $\beta \geq 3/4$, take $p = 1$ and assume (A(1)). If $\beta < 3/4$, take $p = 2$ and assume (A(2)). As a consequence of Lemma 3.4, (33), (36) and Remark 3.3 we obtain (21).

3.2.1 Proof of Theorem 2.4

In Lemma 3.4 we have a bound on the distance between the uniform empirical and the uniform quantile processes, but it does not say anything about its optimality. To obtain this note, that for any $1 \leq p < (2\beta - 1)^{-1}$ we have by (36) and as in (38)

$$\begin{aligned} & \sup_{y \in (0,1)} |\alpha_n(y) - u_n(y) + \sigma_{n,1}^{-1}(\tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y)))| \\ & \leq \sup_{y \in (0,1)} |\alpha_n(y) - \alpha_n(U_n(y)) + \sigma_{n,1}^{-1}(\tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y)))| \quad (44) \\ & \quad + \sup_{y \in (0,1)} |\alpha_n(U_n(y)) - u_n(y)| = O_{\text{a.s.}}(d_{n,p}) + O_{\text{a.s.}}(\sigma_{n,1}^{-1}). \end{aligned}$$

Now, it is sufficient to deal with the process $(\tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y)))$. We approximate this process via several lemmas.

Lemma 3.5 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. Assume (A(p)) and (B). Under the conditions of Theorem 1.1 we have as $n \rightarrow \infty$,*

$$\sup_{y \in (0,1)} \left| \tilde{V}_{n,1}(y) - \tilde{V}_{n,1}(U_n(y)) + \frac{f^{(1)}(Q(y))}{f(Q(y))} \frac{\tilde{V}_{n,p}(y)}{n} \sum_{i=1}^n X_i \right| = O_{\text{a.s.}}(b_n) + O_{\text{a.s.}}(b_{n,p}).$$

Proof. Applying second order Taylor expansion and recalling that $(f \circ Q)^{(1)}(y) = \frac{f^{(1)}(Q(y))}{f(Q(y))}$, one obtains

$$\begin{aligned} & \sup_{y \in (0,1)} \left| (f(Q(y)) - f(Q(U_n(y)))) \sum_{i=1}^n X_i + n^{-1} \frac{f^{(1)}(Q(y))}{f(Q(y))} \tilde{V}_{n,p}(y) \sum_{i=1}^n X_i \right| \\ & \leq \sup_{y \in (0,1)} \left| \frac{f^{(1)}(Q(y))}{f(Q(y))} \sigma_{n,1} n^{-1} \sum_{i=1}^n X_i \left(u_n(y) + \sigma_{n,1}^{-1} \tilde{V}_{n,p}(y) \right) \right| \\ & \quad + \sup_{y \in (0,1)} |(f \circ Q)^{(2)}(y)| \sup_{y \in (0,1)} (y - U_n(y))^2 \left| \sum_{i=1}^n X_i \right| \\ & = O_{\text{a.s.}}(\sigma_{n,1} n^{-1} a_n \sigma_{n,1} (\log \log n)^{1/2}) + O_{\text{a.s.}}(\sigma_{n,1} n^{-1} d_{n,p} \sigma_{n,1} (\log \log n)^{1/2}) \\ & \quad + O_{\text{a.s.}}(\sigma_{n,1}^3 n^{-2} (\log \log n)^{3/2}) = O_{\text{a.s.}}(b_n) + O_{\text{a.s.}}(b_{n,p}). \end{aligned}$$

The above bound follows from (32), (36), (41) and (21) Theorem 2.1.

◻

Lemma 3.6 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. Assume (A(p)) and (B). Under the conditions of Theorem 1.1 we have as $n \rightarrow \infty$,*

$$\sup_{y \in (0,1)} n^{-1} \frac{f^{(1)}(Q(y))}{f(Q(y))} |\tilde{V}_{n,p}(y) - \tilde{V}_{n,1}(y)| \left| \sum_{i=1}^n X_i \right| = O_{\text{a.s.}}(b_n(\log n)^{1/2}).$$

Proof. We have

$$\begin{aligned} & \sup_{y \in (0,1)} n^{-1} \left| \tilde{V}_{n,p}(y) - \tilde{V}_{n,1}(y) \right| \left| \sum_{i=1}^n X_i \right| \\ & \leq \sup_{y \in (0,1)} |f^{(1)}(Q(y))| n^{-1} |Y_{n,2}| \left| \sum_{i=1}^n X_i \right| + O_{\text{a.s.}} \left(n^{-1} \left| \sum_{r=3}^p Y_{n,r} \right| \left| \sum_{i=1}^n X_i \right| \right). \end{aligned}$$

Using (32), (33), we obtain the result. \odot

Similarly to Lemma 3.6, the next result holds true as well.

Lemma 3.7 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. Assume (A(p)) and (B). Under the conditions of Theorem 1.1 we have as $n \rightarrow \infty$,*

$$\sup_{y \in (0,1)} |\tilde{V}_{n,1}(y) - \tilde{V}_{n,1}(U_n(y)) - (\tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y)))| = O_{\text{a.s.}}(b_n(\log n)^{1/2}).$$

From Lemmas 3.5, 3.6, 3.7 we obtain

Corollary 3.8 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$. Assume (A(p)) and (B). Under the conditions of Theorem 1.1 we have as $n \rightarrow \infty$,*

$$\begin{aligned} & \sup_{y \in (0,1)} \left| \tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y)) + n^{-1} \frac{f^{(1)}(Q(y))}{f(Q(y))} \tilde{V}_{n,1}(y) \sum_{i=1}^n X_i \right| \\ & = O_{\text{a.s.}}(b_n(\log n)^{1/2}) + O_{\text{as}}(b_{n,p}). \end{aligned}$$

Recall that $\tilde{R}_n(y) = \alpha_n(y) - u_n(y)$. Then, by (44),

$$\sup_{y \in (0,1)} |\tilde{R}_n(y) + \sigma_{n,1}^{-1}(\tilde{V}_{n,p}(y) - \tilde{V}_{n,p}(U_n(y)))| = O_{\text{a.s.}}(d_{n,p}) + O_{\text{a.s.}}(\sigma_{n,1}^{-1}).$$

Consequently, via Corollary 3.8,

$$\begin{aligned}
& \sup_{y \in (0,1)} \left| \tilde{R}_n(y) - n^{-1} \sigma_{n,1}^{-1} \frac{f^{(1)}(Q(y))}{f(Q(y))} \tilde{V}_{n,1}(y) \sum_{i=1}^n X_i \right| \\
&= O_{\text{a.s.}}(d_{n,p}) + O_{\text{a.s.}}(\sigma_{n,1}^{-1} b_n (\log n)^{1/2}) + O_{\text{a.s.}}(\sigma_{n,1}^{-1} b_{n,p}) + O_{\text{a.s.}}(\sigma_{n,1}^{-1}) \\
&= O_{\text{a.s.}}(d_{n,p}) + O_{\text{a.s.}}(c_n) + O_{\text{a.s.}}(\sigma_{n,1}^{-1}).
\end{aligned}$$

If $\beta \geq 2/3$, then the bound is $O_{\text{a.s.}}(d_{n,2})$ on assuming (A(2)). If $\beta < 2/3$, taking $p = 3$, via Remark 3.3, we obtain the statement (22) of Theorem 2.4.

◊

3.3 Proof of the optimality in Theorem 2.1

If $\beta < \frac{3}{4}$, then the dominating term in Theorem 2.1 is $O_{\text{a.s.}}(a_n)$.

Fix $y = y_0$. Via (36) and as in (24) we obtain

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \sigma_{n,1}^{-1} n (\log \log n)^{-1} \left| u_n(y_0) + \sigma_{n,1}^{-1} f(y_0) \sum_{i=1}^n X_i \right| \\
&= \limsup_{n \rightarrow \infty} n (\sigma_{n,1} \log \log n)^{-1} |u_n(y_0) - \alpha_n(y_0) + (\alpha_n(y_0) + \sigma_{n,1}^{-1} \tilde{V}_{n,p}(y_0))| \\
&= c(\beta, 1) |f^{(1)}(Q(y_0))|.
\end{aligned}$$

Therefore, via (21), for any $y_0 \in (0, 1)$,

$$c(\beta, 1) |f^{(1)}(Q(y_0))| \leq \limsup_{n \rightarrow \infty} \frac{n}{\sigma_{n,1} \log \log n} \sup_{y \in (0,1)} \left| u_n(y) + \frac{\tilde{V}_{n,p}(y)}{\sigma_{n,1}} \right| = O_{\text{a.s.}}(1)$$

which means that the bound is optimal.

◊

3.4 Proof of Theorems 2.2 and 2.5

3.4.1 Properties of the density-quantile function

Note that under an appropriate smoothness of f , (CsR3) is equivalent to

$$(\text{CsR3(i)}) \quad f(Q(y)) \sim y^{\gamma_1} L_1(y^{-1}), \text{ as } y \downarrow 0,$$

$$(\text{CsR3(ii)}) \quad f(Q(1-y)) \sim (1-y)^{\gamma_2} L_2((1-y)^{-1}), \text{ as } y \uparrow 1,$$

for some numbers $\gamma_1, \gamma_2 > 0$ and some slowly varying functions L_1, L_2 . The parameter γ in (CsR3) and γ_1, γ_2 are related as $\gamma = \gamma_1 \wedge \gamma_2$ (see [12]). Let $\gamma_0 = \gamma_1 \vee \gamma_2$. Under (CsR3(i)) and (CsR3(ii)) we have for any $\mu > 0$,

$$\sup_{y \in (0,1)} \frac{(y(1-y))^{\gamma+\mu}}{f(Q(y))} = O(1). \quad (45)$$

Further, note that if $0 < \gamma_1 < 1$ ($0 < \gamma_2 < 1$) then F has bounded support from the left (from the right) (see [21]). Thus, we assume without loss of generality that both γ_1 and γ_2 are not smaller than 1. In this case, for any $\varepsilon > 0$,

$$f(Q(y)) = O(y^{1-\varepsilon}), \quad y \rightarrow 0. \quad (46)$$

Note also, that (CsR3(i)) and (CsR3(ii)) together with $\gamma_0 > 1$ imply that for any $\mu > 0$,

$$\sup_{y \in (0,1)} \frac{|f^{(1)}(Q(y))|}{f(Q(y))} (y(1-y))^\mu = O(1). \quad (47)$$

Further, by [21, p. 116],

$$(f \circ Q)^{(2)}(y) \sim \kappa \frac{(f^{(1)}(Q(y)))^2}{f^3(Q(y))} \quad (48)$$

as $y \rightarrow 0$. The parameter κ is positive if $\gamma_1 > 1$ or $\kappa = 0$ if $\gamma_1 = 1$. A similar consideration applies to the upper tail.

3.4.2 Weighted law of the iterated logarithm

From (32), (36), (46) and $\delta_n^{-1/2} d_{n,p} = O(1)$ if $p \geq 2$ (i.e. $\beta < \frac{3}{4}$) one obtains

Lemma 3.9 *Let $\beta < \frac{3}{4}$. Under the conditions of Theorem 1.1, as $n \rightarrow \infty$,*

$$\sup_{y \in (\delta_n, 1-\delta_n)} \frac{|\alpha_n(y)|}{(y(1-y))^{1/2}} = O_{\text{a.s.}}((\log \log n)^{1/2}).$$

Using now the same argument as in [8, Theorem 2], we obtain a corresponding result for the linear LRD based uniform quantile process.

Lemma 3.10 *Let $\beta < \frac{3}{4}$. Under the conditions of Theorem 1.1, with some $C_0 > 0$, as $n \rightarrow \infty$,*

$$\sup_{y \in (C_0 \delta_n, 1-C_0 \delta_n)} \frac{|u_n(y)|}{(y(1-y))^{1/2}} = O_{\text{a.s.}}((\log \log n)^{1/2}).$$

From Lemma 3.10, by the same argument as in [8, Theorem 3], as $n \rightarrow \infty$,

$$\sup_{y \in (0, \delta_n)} |u_n(y)| = O_{\text{a.s.}}(a_n), \quad (49)$$

provided $\beta < \frac{3}{4}$. Further, via (32), (36) and (46), as $n \rightarrow \infty$, we obtain for arbitrary $\beta \in (1/2, 1)$ and $1 \leq p < (2\beta - 1)^{-1}$,

$$\sup_{y \in (0, \delta_n)} |\alpha_n(y)| = O_{\text{a.s.}}(\delta_n^{1-\varepsilon}(\log \log n)^{1/2}) + O_{\text{a.s.}}(d_{n,p}) = O_{\text{a.s.}}(a_n) + O_{\text{a.s.}}(d_{n,p}).$$

Recall (41). Let $\theta = \theta_n(y)$ be such that $|\theta - y| \leq \sigma_{n,1} n^{-1} |u_n(y)| = O_{\text{a.s.}}(n^{-(\beta-\frac{1}{2})} L_0(n)(\log \log n)^{1/2})$. Arguing as in [8, Theorem 3], uniformly for $y \in (C_0 \delta_n, 1 - C_0 \delta_n)$, as $n \rightarrow \infty$,

$$\frac{y(1-y)}{\theta(1-\theta)} = O_{\text{a.s.}}(1). \quad (50)$$

3.4.3 Proof of Theorem 2.2

First, we need estimates which will replace a part of the proof of Lemma 3.4. All random variables θ below are as in (50).

Lemma 3.11 *Let $p \geq 1$ be an arbitrary integer such that $p < (2\beta - 1)^{-1}$ and assume that $(C(p))$ is fulfilled. Under the conditions of Theorem 2.2, for any $r = 0, \dots, p-1$, as $n \rightarrow \infty$,*

$$\sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} \psi_1(y) |f^{(r)}(Q(y)) - f^{(r)}(Q(U_n(y)))| = O_{\text{a.s.}}(n^{-(\beta-\frac{1}{2})} L_0(n)(\log \log n)^{1/2}).$$

Proof. Let $\beta < \frac{3}{4}$. Take first $\psi_1(y) = (y(1-y))^{\gamma-\frac{1}{2}+\mu}$. Taking a first order Taylor expansion and bearing in mind that $f^{(r+1)}$ are uniformly bounded, we have

$$\psi_1(y) |f^{(r)}(Q(y)) - f^{(r)}(Q(U_n(y)))| = \frac{(\theta(1-\theta))^{\gamma+\mu}}{f(Q(\theta))} \left(\frac{y(1-y)}{\theta(1-\theta)} \right)^{\gamma+\mu} \frac{|y - U_n(y)|}{(y(1-y))^{1/2}}.$$

Further, under the condition $(C(p))$,

$$\begin{aligned} & |f^{(r)}(Q(y)) - f^{(r)}(Q(U_n(y)))| \\ &= \frac{f^{(r+1)}(Q(\theta))}{f(Q(\theta))} (\theta(1-\theta))^{1/2} \left(\frac{y(1-y)}{\theta(1-\theta)} \right)^{1/2} \frac{|y - U_n(y)|}{(y(1-y))^{1/2}}. \end{aligned}$$

Thus, the result follows by Lemma 3.10, (45) and (50).

If $\beta \geq \frac{3}{4}$, assume (C(1)). We use the appropriate form of ψ_1 , (45) and (50).

◊

From Lemma 3.11, and exactly as in the proof of Lemma 3.4, as $n \rightarrow \infty$,

$$\sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} \psi_1(y) |u_n(y) - \alpha_n(y)| = O_{\text{a.s.}}(a_n) + O_{\text{a.s.}}(d_{n,p}).$$

Consequently, by (49) and the comment below it, as $n \rightarrow \infty$, we have for $\beta < \frac{3}{4}$ and $p < (2\beta - 1)^{-1}$,

$$\sup_{y \in (0,1)} \psi_1(y) |u_n(y) - \alpha_n(y)| = O_{\text{a.s.}}(a_n) + O_{\text{a.s.}}(d_{n,p}).$$

The same estimates are valid for $\beta \geq \frac{3}{4}$, since in this case $\psi_1(y) = O(y)$. Consequently, (22) follows.

◊

3.4.4 Proof of Theorem 2.5

First, we show that Lemma 3.5 remains valid when multiplying by $\psi_2(y)$.

From (47), Theorem 2.2 and estimating as in Lemma 3.5, as $n \rightarrow \infty$, we conclude

$$\begin{aligned} \sup_{y \in (0,1)} (y(1-y)) \psi_1(y) \left| \frac{f^{(1)}(Q(y))}{f(Q(y))} \sigma_{n,1} n^{-1} \sum_{i=1}^n X_i \left(u_n(y) + \sigma_{n,1}^{-1} \tilde{V}_{n,p}(y) \right) \right| \\ = O_{\text{a.s.}}(b_n) + O_{\text{a.s.}}(c_n). \end{aligned} \quad (51)$$

In view of (48), for the term in Lemma 3.5 involving $(f \circ Q)^{(2)}(y)$, we estimate

$$\begin{aligned} (y(1-y))^\mu \frac{(f^{(1)}(Q(\theta)))^2}{f^3(Q(\theta))} (y - U_n(y))^2 \left| \sum_{i=1}^n X_i \right| \\ = \left(\frac{f^{(1)}(Q(\theta))}{f^2(Q(\theta))} \theta(1-\theta) \right)^2 \frac{f(Q(\theta))}{(\theta(1-\theta))^{1-\mu}} \left(\frac{y(1-y)}{\theta(1-\theta)} \right)^{1+\mu} \frac{(y - U_n(y))^2}{y(1-y)} \left| \sum_{i=1}^n X_i \right| \\ = O_{\text{a.s.}}(b_n), \end{aligned} \quad (52)$$

uniformly for $y \in (C_0 \delta_n, 1 - C_0 \delta_n)$, on account of (CsR3), (46), (50), Lemma 3.10 and (32). A similar argument yields the same bound for the right tail.

Further, as $n \rightarrow \infty$,

$$\begin{aligned} & \sup_{y \in (0, C_0 \delta_n)} (y(1-y))^{1+\mu} |\tilde{V}_{n,1}(y) - \tilde{V}_{n,1}(U_n(y))| \\ & \leq C_0 \delta_n^{1+\mu} \sup_{y \in (0,1)} |f(Q(y))| \sum_{i=1}^n X_i = O_{\text{a.s.}}(\delta_n^{1+\mu} \sigma_{n,1} (\log \log n)^{1/2}) = o_{\text{a.s.}}(b_n). \end{aligned} \quad (53)$$

and by (47)

$$\begin{aligned} & \sup_{y \in (0, C_0 \delta_n)} (y(1-y))^{1+\mu} \left| \frac{f^{(1)}(Q(y))}{f(Q(y))} \right| \left| n^{-1} \tilde{V}_{n,p} \sum_{i=1}^n X_i \right| \\ & = \delta_n^{1+\mu/2} \sup_{y \in (0, C_0 \delta_n)} (y(1-y))^{\mu/2} \left| \frac{f^{(1)}(Q(y))}{f(Q(y))} \right| O_{\text{a.s.}} \left(\left(\sum_{i=1}^n X_i \right)^2 / n \right) \\ & = O_{\text{a.s.}}(\delta_n^{1+\mu/2} \sigma_{n,1}^2 n^{-1} \log \log n) = O_{\text{a.s.}}(b_n). \end{aligned} \quad (54)$$

The same argument applies to the interval $(1 - C_0 \delta_n, 1)$. Consequently, by (51), (52), (53), (54) and comparing $(y(1-y))^{1+\mu}$ with $(y(1-y))^\mu \psi_1(y)$, the statement of Lemma 3.5 remains true when multiplying by $\psi_2(y)$. The same holds true for Lemmas 3.6, 3.7 and Corollary 3.8. Consequently, Theorem 2.5 is proven.

The optimality of the bound in Theorem 2.2 follows from Theorem 2.5 in the same way we proved optimality in Theorem 2.1.

◉

3.5 Proof of Theorem 2.8

Let $\beta < \frac{3}{4}$. Applying a third order Taylor expansion to $f(Q(y))q_n(y)$, one has

$$\begin{aligned} & |u_n(y) - f(Q(y))q_n(y) + \sigma_{n,1} n^{-1} \frac{f^{(1)}(Q(y))}{2f^2(Q(y))} u_n^2(y)| \\ & = \sigma_{n,1}^2 n^{-2} \frac{f(Q(y))(y(1-y))^{3/2}}{6} Q^{(3)}(\theta) \sigma_{n,1}^{-3} n^3 \frac{|y - U_n(y)|^3}{(y(1-y))^{3/2}}. \end{aligned}$$

We have

$$Q^{(3)}(y) = \frac{f^{(2)}(Q(y))}{f^4(Q(y))} - \frac{3(f^{(1)}(Q(y)))^2}{f^5(Q(y))}.$$

By the same argument as the one leading to (52), it suffices to control the second term. We have

$$\begin{aligned} & (y(1-y))^{1/2} f(Q(y)) \frac{(f^{(1)}(Q(\theta)))^2}{f^5(Q(\theta))} (y(1-y))^{3/2} \\ &= \frac{f(Q(y))}{f(Q(\theta))} \left(\frac{f^{(1)}(Q(\theta))}{f^2(Q(\theta))} \theta(1-\theta) \right)^2 \left(\frac{y(1-y)}{\theta(1-\theta)} \right)^2. \end{aligned}$$

Under (CsR3(i)), (CsR3(ii)), in view of [8, Lemma 1] one has

$$\frac{f(Q(y))}{f(Q(\theta))} \leq \left\{ \frac{y \vee \theta}{y \wedge \theta} \times \frac{1 - (y \wedge \theta)}{1 - (y \vee \theta)} \right\}^\gamma. \quad (55)$$

From this, (41), (50) and Lemma 3.10, as $n \rightarrow \infty$, one concludes

$$\begin{aligned} & \sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} (y(1-y))^{1/2} |u_n(y) - f(Q(y))q_n(y) + \frac{\sigma_{n,1}}{n} \frac{f(Q(y))f^{(1)}(Q(y))}{2f^3(Q(y))} u_n^2(y)| \\ &= O_{\text{a.s.}}(\sigma_{n,1}^2 n^{-2} (\log \log n)^{3/2}). \end{aligned} \quad (56)$$

Next, taking Taylor expansion for $(\tilde{V}_{n,1}(y) - \tilde{V}_{n,1}(U_n(y)))$, one obtains

$$\begin{aligned} & \sigma_{n,1}^{-1} (\tilde{V}_{n,1}(y) - \tilde{V}_{n,1}(U_n(y))) = \\ &= \sigma_{n,1}^{-1} \frac{f^{(1)}(Q(y))}{f(Q(y))} (y - U_n(y)) \sum_{i=1}^n X_i + \sigma_{n,1}^{-1} (f \circ Q)^{(2)}(\theta) (y - U_n(y))^2 \sum_{i=1}^n X_i. \end{aligned}$$

Like in (52), as $n \rightarrow \infty$,

$$\begin{aligned} & \sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} (y(1-y))^\mu \sigma_{n,1}^{-1} (f \circ Q)^{(2)}(\theta) (y - U_n(y))^2 \left| \sum_{i=1}^n X_i \right| \\ &= O_{\text{a.s.}}(\sigma_{n,1}^2 n^{-2} (\log \log n)^{3/2}). \end{aligned}$$

If $\beta \geq \frac{3}{4}$, (56) and (57) remain valid if one replaces the weight functions with $(y(1-y))^2$.

Thus,

$$\begin{aligned} & \sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} \psi_3(y) \left| \alpha_n(y) - f(Q(y))q_n(y) - \sigma_{n,1}^{-1} n^{-1} \frac{f^{(1)}(Q(y))}{2} \left(\sum_{i=1}^n X_i \right)^2 \right| \\ &\leq \text{left hand side of (56)} \\ &\quad + \sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} \psi_3(y) |\tilde{R}_n(y) + \sigma_{n,1}^{-1} (\tilde{V}_{n,1}(y) - \tilde{V}_{n,1}(U_n(y)))| \end{aligned}$$

$$\begin{aligned}
& + \sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} \psi_3(y) \left| \sigma_{n,1}^{-1} (\tilde{V}_{n,1}(y) - \tilde{V}_{n,1}(U_n(y))) \right. \\
& \left. + \sigma_{n,1} n^{-1} \frac{f^{(1)}(Q(y))}{2f^2(Q(y))} u_n^2(y) + \sigma_{n,1}^{-1} n^{-1} \frac{f^{(1)}(Q(y))}{2} \left(\sum_{i=1}^n X_i \right)^2 \right| \\
& = O_{\text{a.s.}}(\sigma_{n,2} n^{-2} (\log \log n)^{3/2}) + O_{\text{a.s.}}(d_{n,p}) + O_{\text{a.s.}}(\sigma_{n,1}^{-1} b_n (\log n)^{1/2}) \\
& + \sigma_{n,1} n^{-1} \sup_{y \in (C_0 \delta_n, 1 - C_0 \delta_n)} \psi_3(y) \left| \frac{f^{(1)}(Q(y))}{2} \right| \left| \frac{u_n(y) + \sigma_{n,1}^{-1} \sum_{i=1}^n X_i}{f(Q(y))} \right|^2 \\
& + O_{\text{a.s.}}(\sigma_{n,1}^2 n^{-2} (\log \log n)^{1/2})
\end{aligned}$$

by (56), (57) and (44) together with Lemmas 3.7, 3.8. Moreover, by (CsR3) and via Theorem 2.2 the bound is of the order $O_{\text{a.s.}}(c_n) + O_{\text{a.s.}}(d_{n,p})$, as $n \rightarrow \infty$.

Further, as $n \rightarrow \infty$,

$$\sup_{y \in (0, C_0 \delta_n)} \psi_3(y) \sigma_{n,1}^{-1} n^{-1} \left(\sum_{i=1}^n X_i \right)^2 = O_{\text{a.s.}}(\delta_n^{1+\mu} \sigma_{n,1} n^{-1} (\log \log n)) = O_{\text{a.s.}}(c_n),$$

and $\sup_{y \in (0, C_0 \delta_n)} \psi_3(y) |\alpha_n(y)| = O_{\text{a.s.}}(c_n)$.

Next, having tail monotonicity assumption (CsR4) we may proceed as in [8]. Let $(k-1)/n < y \leq k/n$. If $U_{k:n} \geq y$, then

$$\sup_{y \in (0, C_0 \delta_n)} \psi_3(y) |f(Q(y)) q_n(y)| \leq \sup_{y \in (0, C_0 \delta_n)} \psi_3(y) |u_n(y)| = O_{\text{a.s.}}(\delta_n^{(1+\mu)}) = O_{\text{a.s.}}(c_n).$$

Further, if $U_{k:n} \leq y$, then

$$\sup_{y \in (0, C_0 \delta_n)} \psi_3(y) |f(Q(y)) q_n(y)| \leq C \sigma_{n,1}^{-1} n \sup_{y \in (0, C_0 \delta_n)} y(y(1-y))^{1+\mu} \log(\delta_n/U_{k:n})$$

for $\gamma_1 = 1$. Now,

$$P(U_{1:n} \leq n^{-2} (\log n)^{-3/2}) \leq \sum_{i=1}^n P(U_i \leq n^{-2} (\log n)^{-3/2}) \leq n^{-1} (\log n)^{-3/2}. \tag{58}$$

Consequently, via the Borel-Cantelli Lemma, as $n \rightarrow \infty$, $U_{k:n}^{-1} = o_{\text{a.s.}}(n^2 (\log n)^{3/2})$. Therefore,

$$\sup_{y \in (0, C_0 \delta_n)} \psi_3(y) |f(Q(y)) q_n(y)| = O_{\text{a.s.}}(c_n) \tag{59}$$

follows for $\gamma_1 = 1$.

◊

3.6 Proof of Proposition 2.9

We follow lines of the proof from [8, Theorem 3]. In view of Lemma 3.10 and the Taylor expansion of $f(Q(y))q_n(y)$, the approximation is valid on $(C_0\delta_n, 1 - C_0\delta_n)$, provided $\beta < \frac{3}{4}$. For $\beta \geq \frac{3}{4}$ it remains true by the choice of $\psi_4(y)$.

Having tail monotonicity assumption (CsR4), let $(k-1)/n < y \leq k/n$. If $U_{k:n} \geq y$, then (cf. (3.13) in [8])

$$\sup_{y \in (0, C_0\delta_n)} \psi_4(y) |f(Q(y))q_n(y)| \leq \sup_{y \in (0, C_0\delta_n)} \psi_4(y) |u_n(y)| = O_{\text{a.s.}}(a_n)$$

from (49) if $\beta < \frac{3}{4}$, and by the choice of $\psi_4(y)$ if $\beta \geq \frac{3}{4}$.

If $U_{k:n} \leq y$ and $\beta \in (\frac{1}{2}, 1)$, then for $\gamma = 1$, as $n \rightarrow \infty$,

$$\sup_{y \in (0, C_0\delta_n)} |f(Q(y))q_n(y)| = O_{\text{a.s.}}(\sigma_{n,1}n^{-1}\ell(n))$$

by (58). Moreover, as in (58), $U_{k:n}^{-1} = o_P(n(\log n)^{3/2})$, as $n \rightarrow \infty$. Therefore, for $\gamma > 1$, as $n \rightarrow \infty$,

$$\sup_{y \in (0, C_0\delta_n)} |f(Q(y))q_n(y)| = O_P(\sigma_{n,1}n^{-1}\ell(n)).$$

◻

References

- [1] Bahadur, R.R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37**, 577–580.
- [2] Csörgő, M. (1983). *Quantile Processes with Statistical Applications*. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.
- [3] Csörgő, M., Csörgő, S., Horváth, L., Mason, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14**, 31–85.
- [4] Csörgő, M., Csörgő, S., Horváth, L., Révész, P. (1985). On weak and strong approximations of the quantile process. In: *Proceedings of the seventh conference on probability theory* (Braşov, 1982), 81–95, VNU Sci. Press, Utrecht.

- [5] Csörgő, S., Haeusler, E. and Mason, D. M. (1991). The quantile-transform–empirical-process approach to limit theorems for sums of order statistics. In: *Sums, trimmed sums and extremes*, 215–267, Progr. Probab., **23**, Birkhäuser, Boston, MA.
- [6] Csörgő, M and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. John Wiley & Sons, Ltd., Chichester.
- [7] Csörgő, M. and Kulik, R. (2006). Weak convergence of Vervaat and Vervaat Error processes of long-range dependent Gaussian sequences. *Preprint*.
- [8] Csörgő, M. and Révész, P. (1978). Strong approximation of the quantile process. *Ann. Statist.* **6**, 882–894.
- [9] Csörgő, M. and Révész, P. (1981). *Strong Approximation in Probability and Statistics*. Academic Press, New York.
- [10] Csörgő, M. and Szyszkowicz, B. (1998). Sequential quantile and Bahadur-Kiefer processes. In: *Handbook of Statist.*, **16** (N. Balakrishnan and C.R. Rao, eds), 631–688. North-Holland, Amsterdam.
- [11] Csörgő, M., Szyszkowicz, B. and Wang, L. (2006). Strong Invariance Principles for Sequential Bahadur-Kiefer and Vervaat Error Processes of Long-Range Dependent Sequences. *Ann. Statist.* **34**, 1016–1044.
- [12] Csörgő, M., Zitikis, R. (2002). On the general Bahadur-Kiefer, quantile, and Vervaat processes: old and new. In: *Limit theorems in probability and statistics*, Vol. I (Balatonlelle, 1999), 389–426, János Bolyai Math. Soc., Budapest.
- [13] Giraitis, L. and Surgailis, D. (1999). Central limit theorem for the empirical process of a linear sequence with long memory. *J. Statist. Plann. Inference* **80**, 81–93.
- [14] Giraitis, L. and Surgailis, D. (2002). The reduction principle for the empirical process of a long memory linear process. *Empirical process techniques for dependent data*, 241–255, Birkhäuser Boston, Boston, MA.
- [15] Ho, H.-C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24**, 992–1024.

- [16] Kiefer, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38**, 1323–1342.
- [17] Kiefer, J. (1970). Deviations between the sample quantile process and the sample df. In: *Nonparametric Techniques in Statistical Inference*, 349–357, M.L. Puri, ed., Cambridge University Press.
- [18] Koul, H.L. and Surgailis, D. (2002). Asymptotic expansion of the empirical process of long memory moving averages. In: *Empirical process techniques for dependent data*, 213–239, Birkhäuser Boston, MA.
- [19] Kulik, R. (2006). Sums of extreme values of subordinated long-range dependent sequences: moving averages with finite variance. *Preprint, submitted*.
- [20] Kulik, R. and Ould Haye, M. (2006). Trimmed sums of long-range dependent moving averages. *Preprint, submitted*.
- [21] Parzen, E. (1979). Nonparametric statistical data modeling. *J. Amer. Statist. Assoc.* **74**, 105–131.
- [22] Surgailis, D. (2002). Stable limits of empirical processes of moving averages with infinite variance. *Stochastic Process. Appl.* **100**, 255–274.
- [23] Surgailis, D. (2004). Stable limits of sums of bounded functions of long-memory moving averages with finite variance. *Bernoulli* **10**, 327–355.
- [24] Wang, Q., Lin, Y.-X. and Gulati, C.M. (2003). Strong approximation for long memory processes with applications. *J. Theor. Probab.* **16**, 377–389.
- [25] Wu, W.B. (2003). Empirical processes of long-memory sequences. *Bernoulli* **9**, 809–831.
- [26] Wu, W.B. (2005). On the Bahadur representation of sample quantiles for dependent sequences. *Ann. Statist.* **4**, 1934–1963.